汎化リスク

名前：松島完忠

学籍番号：t211d070

日付:7/18

# [演習 240] 閾値関数の汎化リスク

1.定数おき、分布から1000点ｍｐデータを生成し、頻度分布のプロット結果を図1に示す。

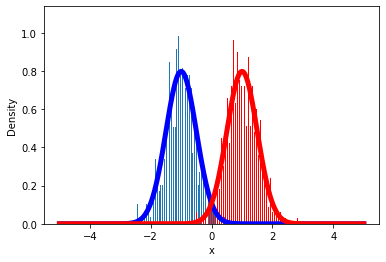


図 1：分布Dｚの頻度分布

2.それぞれに対する閾値関数の汎化リスクを、期待値をプロットした結果を示す。このとき、定数σ＝0.5、1000000点のデータを作成し、それ元に汎化リスクを求めた。四角形

中程度の精度で自動的に生成された説明

図 2：sigma=0.5、経験リスクのプロット

図2より陰性データと陽性データの分布関数が重なる部分で経験リスクが0に近い値となった。

3.1,0.25のときの経験リスクを図3、図4、図5にプロット結果を示す。

アイコン

中程度の精度で自動的に生成された説明

図 3:sigma=2のときの経験リスク

アイコン

自動的に生成された説明

図 4：sigma=1のときの経験リスク

テキスト が含まれている画像

自動的に生成された説明

図 5：sigma=0.25のときの経験リスク

図3、図4，図5から陽性データと陰性データの分布関数が重なる部分が大きいほど経験リスクの最小値は大きくなる。

# 作成プログラム

図6に作成したプログラムを示す。

|  |  |
| --- | --- |
| 1 | import numpy as np |
| 2 | from scipy import stats |
| 3 | import matplotlib.pyplot as plt |
| 4 | import scipy.stats |
| 5 | import matplotlib.ticker as ticker |
| 6 |  |
| 7 | def Genrisk(a,b,theta,N): |
| 8 | sum=0 |
| 9 |  |
| 10 | for (x,y) in zip(a,b): |
| 11 | f=x-theta |
| 12 | if(y\*f<=0): |
| 13 | sum+=1 |
| 14 | return float(sum)/float(N) |
| 15 |  |
| 16 |  |
| 17 | size = 1000000 |
| 18 | sigma = 2 |
| 19 |  |
| 20 | pe=[] |
| 21 | ne=[] |
| 22 | genrisk=[] |
| 23 | x=[] |
| 24 |  |
| 25 | y=stats.uniform.rvs(0,scale=1,size=size) |
| 26 | Y=[] |
| 27 |  |
| 28 |  |
| 29 | for s in y: |
| 30 | if s<0.5: |
| 31 | t=stats.norm.rvs(loc=-1,scale=sigma) |
| 32 | x.append(t) |
| 33 | ne.append(t) |
| 34 | Y.append(-1.0) |
| 35 | else: |
| 36 | t=stats.norm.rvs(loc=1,scale=sigma) |
| 37 | x.append(t) |
| 38 | pe.append(t) |
| 39 | Y.append(1.0) |
| 40 |  |
| 41 |  |
| 42 | xx = np.linspace(-5, 5, 10000) |
| 43 | theta = np.linspace(-5,5,100) |
| 44 |  |
| 45 | for th in theta: |
| 46 | genrisk.append(Genrisk(x,Y,th,size)) |
| 47 |  |
| 48 | fig = plt.figure() |
| 49 | ax=fig.add\_subplot(211,xlabel='x',ylabel='Density') |
| 50 | ax.plot(xx, scipy.stats.norm.pdf(xx,-1,sigma), c='blue',linewidth = 5.0) |
| 51 | ax.plot(xx, scipy.stats.norm.pdf(xx,1,sigma), c='red',linewidth = 5.0) |
| 52 |  |
| 53 |  |
| 54 |  |
| 55 | ax2=fig.add\_subplot(212,xlabel='θ',ylabel='Generalization Risk',ylim=(0,0.6)) |
| 56 | ax2.plot(theta,genrisk,c='black',linewidth=5.0) |

**図6：作成したプログラム**